

A paradox in bosonic energy computations via semidefinite programming relaxations

M. Navascués¹, A. García-Sáez², A. Acín³, S. Pironio⁴, and M. B. Plenio⁵

¹School of Physics, University of Bristol, BS8 1TL, Bristol, U.K.

²Dept. d'Estructura i Constituents de la Matèria, Universitat de Barcelona, 08028 Barcelona, Spain

³ICFO-Institut de Ciències Fotòniques, E-08860 Castelldefels (Barcelona), Spain

⁴Laboratoire d'Information Quantique, Université Libre de Bruxelles, 1050 Bruxelles, Belgium

⁵Universität Ulm, Institut für Theoretische Physik, 89069 Ulm, Germany.

Abstract

We show that the recent hierarchy of semidefinite programming relaxations based on non-commutative polynomial optimization and reduced density matrix variational methods exhibits an interesting paradox when applied to the bosonic case: even though it can be rigorously proven that the hierarchy collapses after the first step, numerical implementations of higher order steps generate a sequence of improving lower bounds that converges to the optimal solution. We analyze this effect and compare it with similar behavior observed in implementations of semidefinite programming relaxations for commutative polynomial minimization. We conclude that the method converges due to the rounding errors occurring during the execution of the numerical program, and show that convergence is lost as soon as computer precision is incremented. We support this conclusion by proving that for any element p of a Weyl algebra which is non-negative in the Schrödinger representation there exists another element \tilde{p} arbitrarily close to p that admits a sum of squares decomposition.

1 Introduction

Computing the energy spectrum of a finite set of indistinguishable particles subject to a given potential is a standard problem appearing in many branches of physics, e.g., in quantum chemistry, atomic physics, or condensed matter physics. Although traditionally the main approaches to this problem have been variational [1], in the last decade it has been attacked with success by means of semidefinite programming (SDP) formulations [2, 3, 4] of the constraints on second-order reduced density matrices proposed in [5, 6, 7] (the so-called 2-RDM method). These SDP methods can be viewed as particular instances of a more general *non-commutative polynomial optimization* approach [8, 9], which extends to the non-commutative setting the method developed by Lasserre [10] and Parrilo [11] for scalar polynomial optimization. Roughly speaking, a non-commutative optimization problem consists in finding the minimal eigenvalue of a hermitian polynomial of non-commutative

operators. To solve such problems, one can define a hierarchy of SDP relaxations, each of which corresponds to finding a sum of squares decomposition of the polynomial to be minimized which provides a lower bound p^k on the optimal solution p^* of the original problem, with $p^1 \leq p^2 \leq \dots \leq p^*$. This approach reduces to the 2-RDM method when applied to fermionic systems, but more generally is also highly successful, e.g., to characterize the set of quantum correlations in quantum information science [12]. Recently, modifications of this algorithm exploiting translational invariance have been proposed independently by Hübener and Barthel [13] and Baumgratz and Plenio [14] as an alternative to variational techniques in condensed matter physics. In [15, 13, 14], it was suggested to apply such SDP methods to compute the ground-state energy of bosonic systems, i.e., to find the minimal eigenvalue of Weyl polynomials.

In this article, we point out that any computer implementation of the SDP hierarchies [8, 9] to bosonic systems will exhibit the non-commutative analog of an effect already observed in similar algorithms for commutative polynomial minimization [16], [17]. On one hand, it can be proven that any relaxation beyond the first one will not provide better lower bounds on p^* , i.e., $p^k = p^1$ for all $k \geq 1$. On the other hand, numerically it is observed that the bounds $\hat{p}^1, \hat{p}^2, \dots$ output by the computer form an increasing sequence, with $\lim_{k \rightarrow \infty} \hat{p}^k = p^*$.

We will show that the resolution of this “mathematical paradox” follows the same lines as the commutative one. Namely, even though there exist positive Weyl polynomials p that do not admit a sum of squares decomposition, for any such polynomial there exists an arbitrarily small perturbation $p \rightarrow \tilde{p} = p + \epsilon g$ such that \tilde{p} can be decomposed as a sum of squares of Weyl polynomials. The rounding errors introduced by the computer while executing the algorithm correspond to such a perturbation, and so numerical implementations of the SDP method converges to the correct answer of the problem. This is, therefore, an example of a numerical method that converges, not in spite of rounding errors, but *because* of them.

This article is structured as follows. In Section 2, we provide the basic definitions and facts about Weyl algebras that are used in the remaining of the text. In Section 3 we will describe the SDP method and illustrate the paradox with a numerical example. The resolution of the paradox will be given in Section 4, where we will prove that any positive Weyl polynomial can be perturbed to a sum of squares of polynomials. Finally, in Section 5 we will present our conclusions.

2 Definitions and basic results on Weyl algebras

2.1 Weyl algebras and Weyl polynomials

A *Weyl algebra* \mathcal{W}_n is a $*$ -algebra with $2n$ generators $\bar{a} = (a_1, a_2, \dots, a_n)$ and $\bar{a}^* = (a_1^*, a_2^*, \dots, a_n^*)$ satisfying the canonical commutation relations (CCRs):

$$[a_i, a_j] = [a_i^*, a_j^*] = 0, [a_i, a_j^*] = \delta_{ij}, \text{ for all } i, j \in \{1, \dots, n\}. \quad (1)$$

An element p of \mathcal{W}_n is thus a linear combination (with complex coefficients) of words in the $2n$ letters \bar{a} and \bar{a}^* . The words and elements of \mathcal{W}_n can also be viewed as monomials and polynomials, respectively, in the $2n$ variables \bar{a} and \bar{a}^* .

Using the CCRs, any element p of a Weyl algebra can be brought to the *normal form*

$$p = \sum_{\bar{s}, \bar{t}} p_{\bar{s}, \bar{t}} (a^*)^{\bar{s}} a^{\bar{t}}, \quad (2)$$

where $\bar{s} = (s_1, \dots, s_n), \bar{t} = (t_1, \dots, t_n) \in \mathbb{N}^n$, $a^{\bar{t}}$ denotes the monomial $a^{\bar{t}} = \prod_{i=1}^n a_i^{t_i}$ and similarly $(a^*)^{\bar{s}} = \prod_{i=1}^n a_i^{*s_i}$, and where $p_{\bar{s}, \bar{t}} \in \mathbb{C}$ are complex coefficients. For instance, the elements $a_1 a_1^*$ and $a_1 a_2 a_2^* a_1^*$, expressed in normal form, are $1 + a_1^* a_1$ and $1 + a_1^* a_1 + a_2 a_2^* + a_1^* a_2^* a_1 a_2$, respectively. We will later show that the decomposition (2) is unique for each $p \in \mathcal{W}_n$. To check whether two different polynomials p, p' in the variables \bar{a}, \bar{a}^* represent the same element $p = p' \in \mathcal{W}_n$ it is therefore enough to write p, p' in normal form.

This last observation suggests a natural norm in \mathcal{W}_n : let $p \in \mathcal{W}_n$, and let (2) be its normal decomposition. Then, we define the l_1 -norm of p as

$$l_1(p) \equiv \sum_{\bar{s}, \bar{t}} |p_{\bar{s}, \bar{t}}|. \quad (3)$$

It is also useful to distinguish the elements of \mathcal{W}_n by resorting to the concept of degree. We say that the degree of an element p of \mathcal{W}_n is $\deg(p) = \max\{\|\bar{s}\|_1 + \|\bar{t}\|_1 : p_{\bar{s}, \bar{t}} \neq 0\}$ for p expressed in normal form (2). It is easy to see that the degree of any monomial $s = \prod_{k=1}^d s_k$ with $s_k \in \{a_j, a_j^*\}_{j=1}^n$ is equal to $\deg(s) \equiv |s| = d$.

Given a monomial s , we denote by $\ddagger s \ddagger$ its *anti-normal ordering*, that is, the monomial s' that results when we reorder the letters appearing in the expression of s in such a way that all the letters a_1, a_2, \dots, a_n end up on the left and all the letters $(a_1^*, a_2^*, \dots, a_n^*)$ end up on the right. For example, $\ddagger a_2^* (a_1^*)^2 a_1 \ddagger = \ddagger a_1^* a_1 a_1^* a_2^* \ddagger = a_1 (a_1^*)^2 a_2^*$.

In this article, we will be mainly concerned with *hermitian* elements of \mathcal{W}_n , i.e., those polynomials p such that $p = p^*$. If p is decomposed as in (2), the hermiticity condition thus translates as $p_{\bar{s}, \bar{t}} = p_{\bar{t}, \bar{s}}^*$, where $p_{\bar{t}, \bar{s}}^*$ denotes the complex conjugate of the number $p_{\bar{t}, \bar{s}}$.

2.2 The Schrödinger representation

Let π be a mapping $\pi : \mathcal{W}_n \rightarrow L(\mathcal{H})$, for some Hilbert space \mathcal{H} , where $L(\mathcal{H})$ denotes the space of linear (not necessarily bounded) operators of \mathcal{H} . We say that π is a $*$ -representation of \mathcal{W}_n if and only if

1. $\pi(1) = \mathbb{I}_{\mathcal{H}}$.
2. $\pi(pq) = \pi(p)\pi(q)$, $\pi(p+q) = \pi(p) + \pi(q)$ for all $p, q \in \mathcal{W}_n$.
3. $\pi(p)^* = \pi(p^*)$, for all $p \in \mathcal{W}_n$.

We now show that \mathcal{W}_n admits a representation. For this, let H be a separable Hilbert space, and let $\{|s\rangle, s \in \mathbb{N}\}$ be an orthonormal basis for H , which we call the *number basis*. If we denote by $\tilde{a} \in L(H)$ the linear operator defined by

$$\tilde{a}|s\rangle = \begin{cases} 0 & \text{for } s = 0, \\ \sqrt{s}|s-1\rangle & \text{otherwise.} \end{cases}$$

then its adjoint \tilde{a}^* satisfies

$$\tilde{a}^*|s\rangle = \sqrt{s+1}|s+1\rangle, \quad (4)$$

and so it can be verified that

$$[\tilde{a}, \tilde{a}^*]|s\rangle = |s\rangle, \text{ for all } s \in \mathbb{N}. \quad (5)$$

Defining $\mathcal{H} = H^{\otimes n}$, we can then build a representation $\pi_S : \mathcal{W}_n \rightarrow L(\mathcal{H})$ for the Weyl algebra \mathcal{W}_n through

$$\pi_S(a_k) = \mathbb{I}^{\otimes k-1} \otimes \tilde{a} \otimes \mathbb{I}^{\otimes n-k}. \quad (6)$$

This representation of \mathcal{W}_n is known as the *Schrödinger (or Fock) representation*. From now on, we always refer to this representation and write π for π_S for simplicity.

2.3 Weyl polynomial minimization

The Schrödinger representation admits a clear physical interpretation: given a set of n one-dimensional particles, it associates to each particle $k \in \{1, \dots, n\}$ a pair of *creation and annihilation operators* $\pi(a_k), \pi(a_k^*)$. The operators describing the position and momentum of particle k along the real line are then given, respectively, by $\pi(x_k) \equiv \pi(\frac{a_k + a_k^*}{\sqrt{2}})$ and $\pi(p_k) \equiv \pi(\frac{a_k - a_k^*}{\sqrt{2}i})$. If these particles are subject to a potential $V(x_1, \dots, x_n)$, the energy operator of the system, in non-relativistic approximation, will be given by

$$\pi(E) = \pi \left(\sum_{i=1}^n \frac{p_i^2}{2m_i} + V(x_1, \dots, x_n) \right), \quad (7)$$

where $m_i \in \mathbb{R}^+$ is the mass of particle i . In particular, the minimum energy of the system will be given by

$$\lambda_{\inf}(E) = \inf \{ \langle \phi | \pi(E) | \phi \rangle : |\phi\rangle \in \mathcal{S}, \langle \phi | \phi \rangle = 1 \}, \quad (8)$$

where \mathcal{S} is the Schwartz space, that is, the set of states $|\phi\rangle$ (i.e., vectors of \mathcal{H}) satisfying $\|\pi(p)\phi\| < \infty$ for all $p \in \mathcal{W}_n$. Note that the minimization over $\langle \phi | E | \phi \rangle$ makes sense, because the energy is an hermitian operator, $E = E^*$, and consequently, $\langle \phi | E | \phi \rangle \in \mathbb{R}$ is a real quantity for all $|\phi\rangle \in \mathcal{S}$.

The case where E is a polynomial in the variable x_i, p_j – or equivalently in the variables a_i, a_j^* – is particularly important (it includes for instance the case where the potential $V(x_1, \dots, x_n)$ is Taylor expanded around some equilibrium position). This motivates the following generic Weyl polynomial optimization problem

$$\lambda_{\inf}(p) = \inf \{ \langle \phi | \pi(p) | \phi \rangle : |\phi\rangle \in \mathcal{S}, \langle \phi | \phi \rangle = 1 \}, \quad (9)$$

for an arbitrary hermitian polynomial $p \in \mathcal{W}_n$. Note that, alternatively, we can write

$$\lambda_{\inf}(p) = \sup \{ \lambda \in \mathbb{R} : \pi(p) - \lambda \geq 0 \}, \quad (10)$$

where positivity is understood in the Schwartz space. This reformulation of the problem will be used in the next section.

2.4 Coherent states

An interesting subset of \mathcal{H} is constituted by the *coherent states*. For any $\alpha \in \mathbb{C}$, denote by $|\alpha\rangle \in H$ the normalized state

$$|\alpha\rangle \equiv e^{-\frac{|\alpha|^2}{2}} \sum_{s \in \mathbb{N}} \frac{\alpha^s}{\sqrt{s!}} |s\rangle. \quad (11)$$

Then, a coherent state in \mathcal{H} is any state of the form $|\bar{\alpha}\rangle = \otimes_{i=1}^n |\alpha_i\rangle$, for any $\bar{\alpha} \in \mathbb{C}^n$. Coherent states are important because they are simultaneous eigenstates of the annihilation operators $\{\pi(a_k) : k = 1, \dots, n\}$. This follows from the easily verified identity $\tilde{a}|\alpha\rangle = \alpha|\alpha\rangle$ for all $\alpha \in \mathbb{C}$. As a result, for any normally-ordered polynomial $p \in \mathcal{W}_n$, we have that

$$\langle \bar{\alpha} | \pi(p) | \bar{\alpha} \rangle = \sum_{\bar{s}, \bar{t}} p_{\bar{s}, \bar{t}} \alpha^{*\bar{s}} \alpha^{\bar{t}} \equiv p(\bar{\alpha}, \bar{\alpha}^*). \quad (12)$$

As an application of the coherent states, let us show that the decomposition (2) is unique, or, equivalently, that 0 admits a unique representation. Suppose thus that

$$0 = \sum_{\bar{s}, \bar{t}} p_{\bar{s}, \bar{t}} a^{*\bar{s}} a^{\bar{t}}, \quad (13)$$

for some coefficients $p_{\bar{s}, \bar{t}} \in \mathbb{C}$. Using the Schrödinger representation, we have that

$$0 = \langle \bar{\alpha} | \pi(0) | \bar{\alpha} \rangle = \sum_{\bar{s}, \bar{t}} p_{\bar{s}, \bar{t}} \alpha^{*\bar{s}} \alpha^{\bar{t}} \quad \text{for all } \bar{\alpha} \in \mathbb{C}^n. \quad (14)$$

The right-hand side is a polynomial in the complex variables $\bar{\alpha}, \bar{\alpha}^*$, and it can only be equal to zero for all values of $\bar{\alpha}$ if $p_{\bar{s}, \bar{t}} = 0$ for all \bar{s}, \bar{t} .

3 SOS decompositions, the SDP hierarchy, and the paradox

Given a hermitian polynomial $p \in \mathcal{W}_n$, we say that p admits a sum-of-squares (SOS) decomposition if there exist polynomials $f_i \in \mathcal{W}_n$ such that

$$p = \sum_i f_i^* f_i. \quad (15)$$

We denote Σ^2 the set of all such polynomials. It is clear that if $p \in \Sigma^2$, then $\pi(p) \geq 0$, since, for any $|\phi\rangle \in \mathcal{S}$,

$$\langle \phi | \pi(p) | \phi \rangle = \sum_i \langle \phi | \pi(f_i^*) \pi(f_i) | \phi \rangle \geq 0. \quad (16)$$

However, the opposite implication is not true, not even in \mathcal{W}_1 . Indeed, as noted by Schmüdgen [18], the family of polynomials $p_\epsilon = (a_1^* a_1 - 1)(a_1 a_1^* - 2) + \epsilon$ satisfies $\pi(p_\epsilon) \geq 0$, for $\epsilon \geq 0$, but nevertheless $p_\epsilon \notin \Sigma^2$, for $\epsilon < \frac{1}{4}$.

Given a $p \in \mathcal{W}_n$, a possible scheme for finding a lower bound λ^* on the solution $\lambda_{\inf}(p)$ of (10) is thus to solve the problem

$$\lambda^* \equiv \sup\{\lambda \in \mathbb{R} : p - \lambda \in \Sigma^2\}. \quad (17)$$

This principle is the one behind the polynomial minimization algorithms developed by Lasserre [10] and Parrilo [11], and their non-commutative analogue [8, 9]. Such algorithms work by searching for SOS decompositions of $p - \lambda$ with some degree constraint. Applied to the Weyl minimization problem, this results in the following sequence of programs:

$$\lambda^k \equiv \sup\{\lambda \in \mathbb{R} : p - \lambda \in \Sigma_k^2\}. \quad (18)$$

Here k is an integer such that $2k \geq \deg(p)$ and indexing the successive programs in the sequence and Σ_k^2 is the set of polynomials which admit a decomposition of the form (15) with $\deg(f_i) \leq k$, for all i . Each of these problems is a semidefinite program, as one can check that

$$\lambda^k = \max\{\lambda \in \mathbb{R} : p - \lambda = (\bar{w}^k)^* Z \bar{w}^k, Z \geq 0\}, \quad (19)$$

where \bar{w}^k is a vector whose components are the normally ordered monomials of degree $\leq k$. Clearly, $\lambda^k \leq \lambda^{k+1}$ and $\lim_{N \rightarrow \infty} \lambda^k = \lambda^*$. The programs (18), (19) thus form a converging hierarchy of SDP relaxations for the problem (17). Supplemented with a boundedness condition (that is not satisfied in the present case of Weyl polynomials), it can be further be shown that this hierarchy necessarily converges to the optimal solution of the problem (10) [8, 9], as problems (17) and (10) then turn out to be equivalent [19].

Let L denote an arbitrary functional on the Weyl algebra, i.e., $L : \mathcal{W}_n \rightarrow \mathbb{C}$. Then, the dual of problems (18), (19) can be shown to be

$$\mu^k = \min\{L(p) : L(1) = 1, L(qq^*) \geq 0 \text{ for all } q \in \mathcal{W}_n, \deg(q) \leq k\}, \quad (20)$$

or explicitly in SDP form

$$\mu^k = \min\left\{\sum_{\bar{s}, \bar{t}} p_{\bar{s}, \bar{t}} y_{\bar{s}, \bar{t}} : y_{\bar{0}, \bar{0}} = 1, M_k(y) \geq 0\right\}, \quad (21)$$

In this last formulation, $p_{\bar{s}, \bar{t}}$ are the coefficients of p in normal form (2), $y \equiv \{y_{\bar{s}, \bar{t}} : \|\bar{s} + \bar{t}\|_1 \leq 2k\} \subset \mathbb{C}$ are the optimization variables¹, and $M_k(y)$ is the *moment matrix of order k* , a matrix whose rows and columns are indexed by pairs of vectors $(\bar{s}, \bar{t}) \in \mathbb{N}^n \times \mathbb{N}^n$, with $\|\bar{s} + \bar{t}\|_1 \leq k$, and with entries defined by

$$[M_k(y)]_{(\bar{s}, \bar{t}), (\bar{u}, \bar{v})} = \sum_{\bar{q}, \bar{r}} c_{\bar{q}, \bar{r}} y_{\bar{q}, \bar{r}}, \quad (22)$$

where $\{c_{\bar{q}, \bar{r}}\}$ are the normal form coefficients of the monomial $(a^{*\bar{s}} a^{\bar{t}})^* a^{*\bar{u}} a^{\bar{v}}$, i.e.,

$$(a^{*\bar{s}} a^{\bar{t}})^* a^{*\bar{u}} a^{\bar{v}} = \sum_{\bar{q}, \bar{r}} c_{\bar{q}, \bar{r}} a^{*\bar{q}} a^{\bar{r}}. \quad (23)$$

The next lemma shows that problems (18), (19) and (20), (21) are, in fact, equivalent:

¹Note that, if $p_{\bar{s}, \bar{t}} \in \mathbb{R}, \forall \bar{s}, \bar{t}$, one can assume $\{y_{\bar{s}, \bar{t}} : \|\bar{s} + \bar{t}\|_1 \leq 2k\} \subset \mathbb{R}$.

Lemma 1. *If there exists a feasible point of program (19), then $\mu^k = \lambda^k$.*

Proof. By Sylvester's criterion [20], it just suffices to show that problem (20) admits a strictly feasible point, i.e., that there exists a functional L such that $L(1) = 1$ and $L(pp^*) > 0$, for all $p \neq 0$. Now, consider the functional $L(p) = \text{tr}(\Omega \pi(p))$, with

$$\Omega = \frac{1}{(2\pi)^n} \int d\bar{\alpha} e^{-|\bar{\alpha}|^2} |\bar{\alpha}\rangle\langle\bar{\alpha}|. \quad (24)$$

This functional satisfies $L(1) = 1$. Also, for any non-zero polynomial $p \neq 0 \in \mathcal{W}_n$ we have that

$$\begin{aligned} L(p^*p) &= \frac{1}{(2\pi)^n} \int d\bar{\alpha} e^{-|\bar{\alpha}|^2} \langle\bar{\alpha}|\pi(p^*p)|\bar{\alpha}\rangle \geq \frac{1}{(2\pi)^n} \int d\bar{\alpha} e^{-|\bar{\alpha}|^2} |\langle\bar{\alpha}|\pi(p)|\bar{\alpha}\rangle|^2 \\ &= \frac{1}{(2\pi)^n} \int d\bar{\alpha} e^{-|\bar{\alpha}|^2} |p(\bar{\alpha}, \bar{\alpha}^*)|^2 > 0. \end{aligned} \quad (25)$$

The last inequality comes from the fact that $p(\bar{\alpha}, \bar{\alpha}') \neq 0$ and that the integration takes place in all \mathbb{C}^n . \square

Even though the above SDP hierarchy cannot be guaranteed to converge to the optimal value $\lambda_{\inf}(p)$ of (10), every SDP step provides a lower-bound on $\lambda_{\inf}(p)$. Based on the successful applications of this SDP hierarchy to fermionic systems and quantum correlations, where in practice very good lower bounds are obtained after only a few SDP relaxations, one could expect a good overall performance also in the context of Weyl polynomials. However, as the next result shows, no improvement over the first lower-bound can be obtained by considering higher steps in the hierarchy.

Lemma 2. *Let $p \in \Sigma^2$. Then, $p \in \Sigma_k^2$ for $2k \leq \deg(p)$.*

Proof. By hypothesis, $p \in \Sigma^2$, and thus there exist polynomials f_i such that (4) holds. We will show that all such polynomials satisfy $2\deg(f_i) \leq \deg(p)$.

Suppose, on the contrary, that

$$p = \sum f_i^* f_i + f^* f, \quad (26)$$

with $2\deg(f) > \deg(p)$. Then,

$$p(\bar{\alpha}, \bar{\alpha}^*) = \langle\bar{\alpha}|\pi(p)|\bar{\alpha}\rangle \geq \langle\bar{\alpha}|\pi(f^*)\pi(f)|\bar{\alpha}\rangle \geq |\langle\bar{\alpha}|\pi(f)|\bar{\alpha}\rangle|^2 = |f(\bar{\alpha}, \bar{\alpha}^*)|^2. \quad (27)$$

Now, denote by $LT(f)$ the leading terms of f , i.e.,

$$LT(f)(\bar{\alpha}, \bar{\alpha}^*) = \sum_{\|\bar{s}+\bar{t}\|_1=\deg(f)} f_{\bar{s},\bar{t}}(\bar{\alpha}^{\bar{s}})^* \bar{\alpha}^{\bar{t}}, \quad (28)$$

and choose $\bar{\beta} \in \mathbb{C}^n$ such that $LT(f)(\bar{\beta}, \bar{\beta}^*) = c \neq 0$. Then it is straightforward that $|f(r\bar{\beta}, r\bar{\beta}^*)| = O(r^{\deg(f)})$. However, $|p(r\bar{\beta}, r\bar{\beta}^*)| \leq O(r^{\deg(p)})$. For (27) to hold, we must thus have that $2\deg(f) \leq \deg(p)$. \square

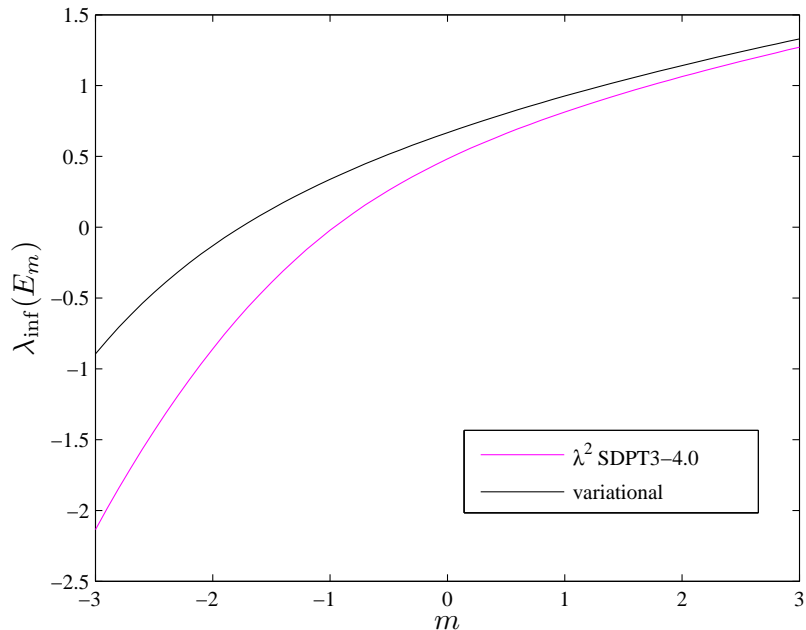


Figure 1: Bounds on $\lambda_{\inf}(E_m)$ as a function of m . The upper curve represents an upper bound obtained through variational techniques. The lower curve represents the lower bound λ^2 corresponding to the first relaxation of the problem.

What Lemma 2 shows is that, for any polynomial p , the sequence of values $\lambda^{k_0}, \lambda^{k_0+1}, \dots$, with $k_0 = \lceil \deg(p)/2 \rceil$ is constant and equal to λ^* . In other words: the first SDP relaxation of the problem already provides the best approximation to $\lambda_{\inf}(p)$ attainable with SOS decompositions.

How does such an approximation perform? Consider the uniparametric family of one-dimensional hamiltonians $\{E_m : m \in \mathbb{R}\}$, with

$$E_m = \frac{1}{2}\tilde{p}^2 + m\tilde{x}^2 + \tilde{x}^4, \quad (29)$$

with $\tilde{x} \equiv (\tilde{a} + \tilde{a}^*)/\sqrt{2}$, $\tilde{p} \equiv (\tilde{a} - \tilde{a}^*)/i\sqrt{2}$. Note that, for $m < 0$, E_m corresponds to the interesting double-well potential. Figure 1 shows a plot of λ^2 as a function of m , together with an upper bound on E_m obtained through variational methods. We used the solver SDPT3-4.0 [21] and the MATLAB package YALMIP [22] to carry out the SDP calculations. It is clear that, as soon as $m < 0$, the approximation given by λ^2 becomes worse.

From our discussions above, it follows that $\lambda^2 = \lambda^3 = \lambda^4 = \dots$, and so λ^2 represents the best lower bound on $\lambda_{\inf}(e_m)$ achievable with the SDP hierarchy. Figure 2 shows, however, that such is not the case. Indeed, we see that subsequent relaxations of the problem return lower bounds which are closer and closer to the variational upper bound, until, at λ^6 , both bounds become practically indistinguishable. What is happening?

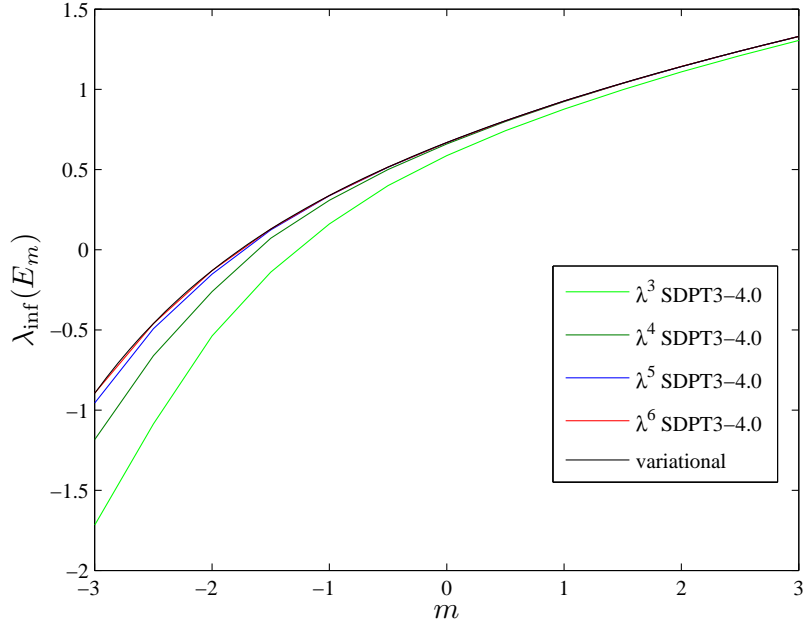


Figure 2: Lower-bounds on $\lambda_{\inf}(E_m)$ as a function of m corresponding to the second, third, fourth and fifth SDP relaxations, in ascending order. The black line represents the variational estimate, quite close to λ^6 (red line).

4 Resolution of the paradox

As mentioned in the introduction, the above paradox is not new in commutative polynomial optimization: indeed, Henrion and Lasserre [16] noticed that the numerical implementation of their SDP algorithm for polynomial minimization returned the optimal value of the 2-dimensional Motzkin polynomial, instead of its SOS value ($-\infty$). Lasserre successfully solved this paradox by proving that any commutative positive polynomial can be approximated arbitrarily well by an SOS decomposition [23]. The accepted resolution of the paradox was that the rounding errors occurring during the numerical computations perturbed the polynomial p to be minimized to another one of higher degree admitting an SOS decomposition [17].

In this Section we will prove a non-commutative analog of this result, namely, that Weyl polynomials which are positive semidefinite in the Schrödinger representation can be perturbed to a higher degree polynomial in Σ^2 . This is formally stated in the following Theorem:

Theorem 3. *Let p be an element of \mathcal{W}_n such that $\pi(p) \geq 0$. Then, for any $\epsilon, \delta > 0$, there exists a polynomial $\tilde{p} \in \Sigma^2$ such that $l_1(p - \tilde{p}) < \epsilon$ and $|\lambda_{\inf}(\tilde{p}) - \lambda_{\inf}(p)| < \delta$.*

The proof of this theorem follows straightforwardly from the next three lemmas. In these lemmas, the constant c is arbitrary but fixed to be $c > 2$.

Lemma 4. Let $p \in \mathcal{W}_n$ be a polynomial such that $\pi(p) \geq 0$. Let

$$g_c^r \equiv \sum_{\|\bar{t}\|_1 \leq r} \frac{(n-1)!}{c^{\|\bar{t}\|_1} (n + \|\bar{t}\|_1 - 1)!} a^{\bar{t}} (a^{\bar{t}})^*, \quad (30)$$

where the sum runs over all vectors $\bar{t} \in \mathbb{N}^n$ of length $\|\bar{t}\|_1$ less or equal than r . Then for any $\epsilon > 0$, there exists some r_0 such that for all $r > r_0$,

$$\tilde{p}^r = p + \epsilon g_c^r \in \Sigma^2. \quad (31)$$

Lemma 5. Let \tilde{p}^r be defined as in Lemma 4. Then

$$l_1(p - \tilde{p}^r) \leq \epsilon \cdot \frac{c}{c-2}, \quad (32)$$

for any $r > 0$.

Lemma 6. Let p be a hermitian polynomial in \mathcal{W}_n such that $\lambda_{\inf}(p)$ exists, and let g_c^r be defined as in (30). Then, for any $\delta > 0$, there exists a number $q \in \mathbb{R}^+$ such that

$$\lambda_{\inf}(p) \leq \lambda_{\inf}(p + \epsilon g_c^r) \leq \lambda_{\inf}(p) + \delta + \epsilon q, \quad (33)$$

for all $r, \epsilon > 0$.

The proofs of Lemma 4 and 6 are given here below. The proof of Lemma 5 follows from the results presented in Appendix A. This Lemma implies that p and \tilde{p}^r can be made arbitrarily close if we take ϵ small enough. Moreover, if we express $p - \tilde{p}^r$ in the anti-normal form, i.e., $p - \tilde{p}^r = \sum_{\bar{s}, \bar{t}} p_{\bar{s}, \bar{t}}^a a^{\bar{s}} (a^{\bar{t}})^*$, the corresponding natural norm $l_1^a(p - \tilde{p}^r) \equiv \sum_{\bar{s}, \bar{t}} |p_{\bar{s}, \bar{t}}^a|$ is trivially bounded by $\epsilon \cdot e^{1/c}$. This implies that any computer implementation of program (19) where the polynomial to minimize is expressed in normal or anti-normal form will require a lot of precision in order to distinguish \tilde{p}^r from p for low values of ϵ .

Proof of Lemma 4. The demonstration of this Lemma will make use of two lemmas, proven in Appendices B, C, respectively.

Lemma 7. Let $s \in \mathcal{W}_n$ be a monomial. Then,

$$\dagger s s^* \dagger - s s^* \in \Sigma^2. \quad (34)$$

Lemma 8. Let L be a linear functional in \mathcal{W}_n . If $L(h^* h) \geq 0$ for any $h \in \mathcal{W}_n$ and there exist $c, d > 0, k \in \mathbb{N}$ such that for any monomial $s \in \mathcal{W}_n$, the relation

$$|L(s)| \leq d c^{|s|} \Gamma\left(\frac{|s| + k + 1}{2}\right) \quad (35)$$

holds, then there exists a non normalized quantum state ρ (i.e., a non-negative, trace-class operator) such that $L(h) = \text{tr}(\rho \pi(h))$, for any polynomial $h \in \mathcal{W}_n$.

Let us now proceed with the proof of Lemma 4. Following Lasserre et al. [24], let $\mathcal{W}_n(r)$ be the set of elements of \mathcal{W}_n with degree less or equal than r , and consider the semidefinite program:

$$\epsilon_r^* = \min_L \{L(p) | L : \mathcal{W}_n(2r) \rightarrow \mathbb{R}, L \text{ linear}, L(g_c^r) \leq 1, L(h^*h) \geq 0, \forall h \in \mathcal{W}_n(r)\}. \quad (36)$$

Noting that $L = 0$ is an admissible linear functional, we have that the problem has feasible points and that $\epsilon_r^* \leq 0$. Condition $L(g_c^r) \leq 1$, together with Lemma 7, implies that the diagonal entries of all feasible moment matrices $M_r(y)$ are upper bounded, and so the absolute values of the rest of the entries, due to positive semidefiniteness. From these two observations, it follows that our problem admits a solution, i.e., $\epsilon_r^* \neq -\infty$ is attainable for a feasible choice of L .

The dual of (36) is

$$\max_{\epsilon} \{-\epsilon : \epsilon \geq 0, p + \epsilon g_c^r \in \Sigma^2\}. \quad (37)$$

That this problem has solutions for any $p = p^*$ follows from the fact that, for all $\mu \in \mathbb{C}$ and any pair of monomials $s, t \in \mathcal{W}_n$,

$$\mu s^*t + \mu^*t^*s + |\mu|\{s^*s + t^*t\} \in \Sigma^2. \quad (38)$$

And thus, invoking Lemma 7,

$$\mu s^*t + \mu^*t^*s + |\mu|\{\dagger s^*s \dagger + \dagger t^*t \dagger\} \in \Sigma^2. \quad (39)$$

By increasing the value of ϵ , at some point we will therefore have that $p + \epsilon g_c^r \in \Sigma^2$.

Moreover, in this particular case, there is no duality gap, i.e., the solutions of both the primal and dual problems coincide. Again, this can be established by invoking the quantum state (24): choosing $\sigma > 0$ such that $\text{tr}\{\Omega\pi(g_c^r)\} < 1/\sigma$, it follows that $L(h) \equiv \sigma \text{tr}(\Omega\pi(h))$ is a strictly feasible point of (36) and, thus, the solutions of the dual and primal problems are the same [20].

This, together with the fact that $g_c^s - g_c^r$ is a sum of squares for $s \geq r$, implies that, for all $\epsilon \geq -\epsilon_r^* \geq 0$, the polynomial

$$p + \epsilon g_c^s \quad (s \geq r) \quad (40)$$

is also a sum of squares.

The sequence $(\epsilon_r^*)_r$ is, therefore, and increasing one. We will next proof that $\lim_{r \rightarrow \infty} \epsilon_r^* = 0$, and so that the $\epsilon > 0$ appearing in the formulation of Lemma 4 can be taken arbitrarily small.

Consider the sequence L_r^* of functionals that attain the solutions ϵ_r^* of the problem, and denote by M_r^* their corresponding moment matrices (the entries $(M_r^*)_{(\bar{s}, \bar{t}), (\bar{u}, \bar{v})}$ where either $\|\bar{s} + \bar{t}\|_1 > r$ or $\|\bar{u} + \bar{v}\|_1 > r$ are assumed to be completed with zeros). By Lemma 7, we have that

$$(M_r^*)_{(\bar{s}, \bar{t}), (\bar{s}, \bar{t})} \leq (M_r^*)_{(\bar{s} + \bar{t}, 0), (\bar{s} + \bar{t}, 0)} \leq \frac{(n + \|\bar{s} + \bar{t}\|_1 - 1)!}{(n - 1)!} c^{\|\bar{s} + \bar{t}\|_1} =: d(\bar{s}, \bar{t}). \quad (41)$$

Now, perform the transformation

$$(M_r^\star)_{(\bar{s}, \bar{t}), (\bar{u}, \bar{v})} \rightarrow (N_r^\star)_{(\bar{s}, \bar{t}), (\bar{u}, \bar{v})} \equiv d(\bar{s}, \bar{t})^{1/2} (M_r^\star)_{(\bar{s}, \bar{t}), (\bar{u}, \bar{v})} d(\bar{u}, \bar{v})^{1/2}. \quad (42)$$

N_r^\star is thus positive semidefinite and its diagonals are upper bounded by 1 for all r ; it follows that all the entries of the matrices N_r^\star are in the interval $[1, -1]$. By the Banach-Alaoglu theorem, the sequence N_r^\star admits a subsequence $\{r_i\}$ that converges in the weak-* topology to a limit $N_{r_i}^\star \rightarrow \hat{N}$ when $i \rightarrow \infty$ [25]. Undoing the previous change of coordinates, we are left with an infinite sized matrix \hat{M} that defines a linear functional $\hat{L}(p) \equiv \sum_{\bar{s}, \bar{t}} p_{(\bar{s}, \bar{t})} \hat{M}_{(0, \bar{s}), (0, \bar{t})}$ on the Weyl algebra.

This functional satisfies $\hat{L}(h^*h) \geq 0$, for any polynomial h . Moreover, for any sequence $s = s_1 s_2$, with $|s_1| = \left\lceil \frac{|s|}{2} \right\rceil$, $|s_2| = \left\lfloor \frac{|s|}{2} \right\rfloor$,

$$\begin{aligned} |\hat{L}(s)| &\leq \left(\hat{L}(s_1 s_1^*) \hat{L}(s_2^* s_2) \right)^{1/2} \leq \left(\frac{(n + |s_1| - 1)! (n + |s_2| - 1)!}{(n - 1)! (n - 1)!} \right)^{1/2} \sqrt{c}^{|s_1| + |s_2|} \\ &\leq \frac{\sqrt{c}^{|s|}}{(n - 1)!} \Gamma \left(\frac{2n + 2 + |s|}{2} \right). \end{aligned} \quad (43)$$

By Lemma 8, this last condition implies that there exists a non normalized quantum state $\rho \in S_1(\mathcal{H})$ in the Schrödinger representation such that $\hat{L}(h) = \text{tr}(\rho \pi(h))$, for all h .

Now,

$$\lim_{r \rightarrow \infty} \epsilon_r^\star = \hat{L}(p) = \text{tr}(\rho \pi(p)) \geq 0, \quad (44)$$

where the last inequality follows from the non-negativity assumption on p . On the other hand, $\epsilon_r^\star \leq 0 \forall r$, and, therefore, we have that $\lim_{r \rightarrow \infty} \epsilon_r^\star = 0$. □

Proof of Lemma 6. Given a vector $\bar{m} \in \mathbb{N}^n$, we will denote by $|\bar{m}\rangle$ the number state $|m_1\rangle \otimes |m_2\rangle \otimes \dots \otimes |m_n\rangle$. Now, if $\pi(p) \geq 0$, then $\lambda_{\inf}(p)$ exists and can be written as the limit of a sequence of the form $\langle \phi_i | \pi(p) | \phi_i \rangle$, where $\{\phi_i\}$ are normalized quantum states. Such states can be, in turn, approximated with arbitrary precision by finite linear combinations of number states. Choose, then, a number $M \in \mathbb{N}$ such that the normalized state $|\Phi\rangle \equiv \sum_{\|\bar{m}\|_\infty \leq M} d_{\bar{m}} |\bar{m}\rangle$ satisfies $\lambda_{\inf}(p) \leq \langle \Phi | \pi(p) | \Phi \rangle \leq \lambda_{\inf}(p) + \delta$.

It can be verified that, for any pair of number states $|\bar{m}\rangle, |\bar{m}'\rangle$, with $\|\bar{m}\|_\infty, \|\bar{m}'\|_\infty \leq M$, and any monomial s of the annihilation and creation operators, the inequality $|\langle \bar{m} | s s^* | \bar{m}' \rangle| \leq \frac{(|s| + M)!}{M!} \delta_{\bar{m}, \bar{m}'}$ holds. Therefore,

$$\langle \Phi | s s^* | \Phi \rangle \leq \frac{(|s| + M)!}{M!} \cdot \sum_{\bar{m}} |d_{\bar{m}}|^2 = \frac{(|s| + M)!}{M!}. \quad (45)$$

Finally, choose $c > 1$. It follows that

$$\langle \Phi | g_c^r | \Phi \rangle \leq \sum_{l=0}^{\infty} \frac{\#\{\bar{t} : |\bar{t}| \leq l\}}{\frac{(n+l-1)!}{(n-1)!l!}} \frac{(l + M)!}{M!l!c^l} = \sum_{l=0}^{\infty} \frac{(l + M)!}{M!l!c^l} = \left(\frac{c}{c - 1} \right)^{M+1}, \quad (46)$$

where in order to identify the second and third expressions we made use of Proposition 11 in Appendix A. Equating to q the last result, we arrive at the promised Lemma. \square

Note that Lemma 8 alone provides an alternative explanation for the observed convergence to the optimal solution, this time from the point of view of the dual problem (20). The reason why μ^k does not necessarily converge in theory to $\lambda_{\inf}(p)$ for the exact problem is because the values $L(s)$ in each program grow faster than $dc^{|s|}\Gamma(\frac{|s|+l}{2})$. If, however, due to finite numerical precision our solvers limit the magnitude of such momenta, low order relaxations μ^k should provide a better approximation to λ_{\inf} ².

Increasing the numerical precision of our programs should thus have a two-fold effect: on one hand, it should allow the computer to distinguish between p and its perturbation \tilde{p} . On the other hand, it should extend the moment matrix search space to include matrices with entries of very different magnitude. A high precision numerical computation should therefore make the curves in Figure 2 collapse to the same line.

We used the semidefinite programming solver SDPA-GMP [4, 26] to compute SDP approximations of E_{-1}, E_1 with a precision of 600 digits. Figures 3 and 4 show the outputs $\lambda^k - \lambda^2$ of both problems as a function of *lambdaStar*, an internal parameter of SDPA-GMP that constrains the magnitudes of the entries of the moment matrix³ [26]. Notice that, in agreement with the above interpretation, the difference between λ^k and $\lambda^2 = \lambda^*$ tends to zero as the constraints on the moment matrix disappear (right end of Figures 3 and 4). Conversely, the solutions of higher order relaxations start to differ from λ^2 and become closer to the actual solution of the problem as we restrict the magnitude of the entries of the moment matrix (left end of Figures 3 and 4).

5 Conclusion

We have identified a paradoxical behaviour in the application to bosonic systems of the SDP methods widely used in quantum chemistry energy calculations and quantum information. Namely, we have pointed out that numerical implementations of the method seems to converge despite a simple theoretical argument showing that the first SDP relaxation should already provide the best lower-bound to the problem at hand. This phenomenon is similar to an analogous behavior observed in commutative polynomial optimization [16, 17] and we suggested that the paradox arises from rounding errors introduced in the numerical computation. We provided a theoretical basis for this assumption by proving that for any bosonic hamiltonian to be minimized there exists a perturbation of it whose ground state energy can accurately be approximated by the SDP method. Furthermore, we showed that the effect disappears as soon as we increase the computer precision.

Our results suggest that the above problem could be avoided by constraining the values of the diagonal elements of the moment matrices in each program. Thanks to such constraints,

²This does not apply to high order relaxations corresponding to $k \gg 1$, as a constraint of the form $L(a_j^k(a_j^k)^*) < K$ becomes unimplementable as soon as $K > k!$.

³More concretely, *lambdaStar* is such that $(\text{lambdaStar}) \cdot \mathbb{I} - M_k(y) \geq 0$.

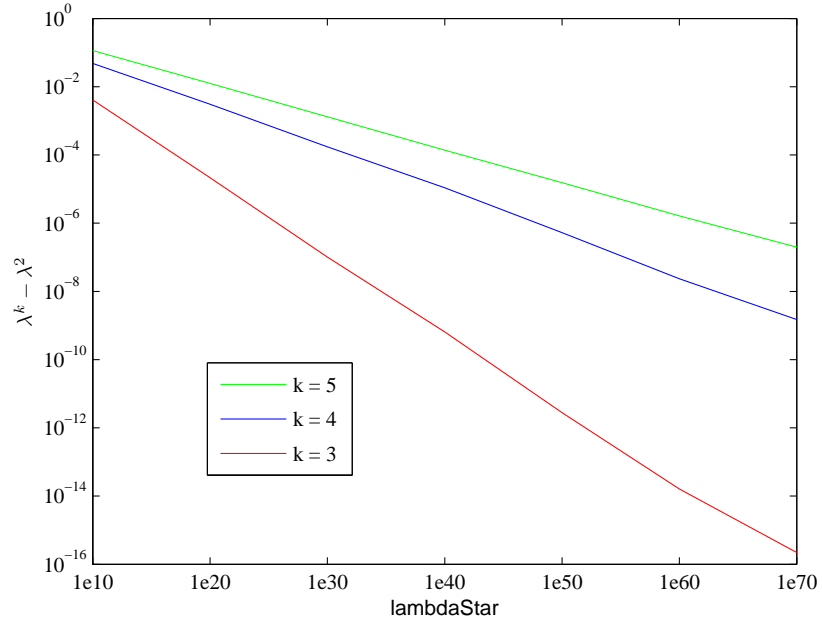


Figure 3: Plot of $\lambda^k - \lambda^2$ as a function of the parameter λ_{star} in the case $m = -1$.

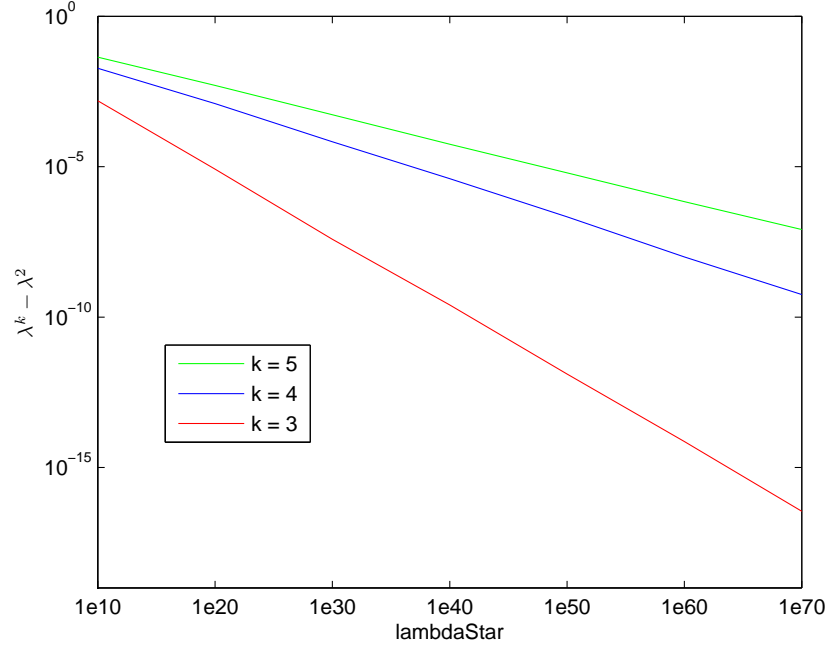


Figure 4: Plot of $\lambda^k - \lambda^2$ as a function of the parameter λ_{star} in the case $m = 1$.

computer implementations of the primal problem could return reliable solutions without having to resort to extremely high precision numerical calculations. This approach will be explored in a forthcoming article.

It is worth noting that Cimprič [27] proposed to use Schmüdgen’s positivstellensatz for Weyl algebras [18] to introduce a different SDP hierarchy than the one presented here in order to find rigorous lower bounds on the minimum value of arbitrary Weyl polynomials. The application of this method, however, requires high precision SDP solvers.

Let us conclude with a problem for the Noncommutative Real Algebraic Geometry community. We have shown that the set of SOS polynomials is dense in the set of positive semidefinite elements of the Weyl algebra, i.e., we can approximate any polynomial which is positive semidefinite in the Schrödinger representation by a SOS. It would be interesting to know if this kind of results also hold in other algebras important for quantum chemistry. For instance, if such an ‘approximation property’ were also true in algebras containing coulombian elements of the type $1/|\bar{x}_i - \bar{x}_j|$, then we would be able to estimate electronic molecular energies without the need of introducing orbital basis sets.

6 Acknowledgements

M. N. has been supported by the Templeton Foundation. We acknowledge Monique Laurent and Frank Vallentin for interesting discussions. S.P. acknowledges financial support from the Brussels-Capital Region through a BB2B grant. MBP has been supported by the Alexander von Humboldt Foundation.

References

- [1] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: introduction to advanced electronic structure theory*, Dover publications Inc., Mineola, New York (1996).
- [2] M. Nakata, H. Nakatsuji, M. Ehara, M. Fukuda, K. Nakata, and K. Fujisawa, J. Chem. Phys., **114**, 8282 (2001).
- [3] D. A. Mazziotti, Phys. Rev. Lett. **93**, 213001 (2004).
- [4] M. Nakata, B. J. Braams, K. Fujisawa, M. Fukuda, J. K. Percus, M. Yamashita, and Z. Zhao, J. Chem. Phys. **128**, 164113 (2008).
- [5] A. J. Coleman, Rev. Mod. Phys. **35**, 668 (1963).
- [6] C. Garrod and J. K. Percus, J. Math. Phys. **5**, 1756 (1964).
- [7] R. M. Erdahl, Intl. J. Quantum Chem. **13**, 697 (1978).
- [8] S. Pironio, M. Navascués and A. Acín, SIAM J. Optim. **20**, 5, 2157-2180 (2010).

- [9] M. Navascués, S. Pironio and A. Acín, *Noncommutative Polynomial Optimization*, Handbook on Semidefinite, Cone and Polynomial Optimization, M.F. Anjos and J. Lasserre (eds); Springer, 2011.
- [10] J. B. Lasserre, SIAM J. Optim., 11, 796-817 (2001).
- [11] P.A. Parrilo, Math. Program. Ser. B 96, 293-320 (2003).
- [12] M. Navascués, S. Pironio and A. Acín, Phys. Rev. Lett. 98, 010401 (2007); M. Navascués, S. Pironio and A. Acín, New J. Phys. 10, 073013 (2008); A. C. Doherty, Y. C. Liang, B. Toner and S. Wehner, Proceedings of IEEE Conference on Computational Complexity 2008, College Park, Maryland, USA, IEEE Computer Society (2009).
- [13] R. Hübener and T. Barthel, arXiv:1106.4966v1.
- [14] T. Baumgratz and M. B. Plenio, New J. Phys. **14**, 023027 (2012).
- [15] J. R. Hammond and D. A. Mazziotti, Phys. Rev. A 71, 062503 (2005).
- [16] D. Henrion and J. B. Lasserre, *Detecting global optimality and extracting solutions in GloptiPoly*, Positive polynomials in control, D. Henrion, and A. Garulli, eds., Lecture Notes on Control and Information Sciences, Vol. 312, Springer Verlag, Berlin (2005).
- [17] H. Waki, M. Nakata and M. Muramatsu, *Strange Behaviors of Interior-point Methods for Solving Semidefinite Programming Problems in Polynomial Optimization*, To appear in Computational Optimization and Applications.
- [18] K. Schmüdgen, Math. Annalen, 331, pp. 779-794 (2005).
- [19] J. W. Helton and S. A. McCullough, *A Positivstellensatz for non-commutative polynomials*, Trans. Amer. Math. Soc., 356 (2004), pp. 3721–3737.
- [20] L. Vandenberghe and S. Boyd, *Semidefinite programming*, S. SIAM Review **38**, 49 (1996).
- [21] K.C. Toh, M.J. Todd and R.H. Tutuncu, *SDPT3 — a Matlab software package for semidefinite programming*, *Optimization Methods and Software*, 11, 545 (1999); R.H. Tutuncu, K.C. Toh and M.J. Todd, *Solving semidefinite-quadratic-linear programs using SDPT3*, Mathematical Programming Ser. B, 95, 189 (2003).
- [22] J. Löfberg, *YALMIP : A Toolbox for Modeling and Optimization in MATLAB*, <http://control.ee.ethz.ch/~joloef/yalmip.php>.
- [23] J. B. Lasserre, SIAM Journal on Optimization, 16, 751 (2006).
- [24] J. B. Lasserre, T. Netzer, Math. Z., 256, 99 (2007).
- [25] M. Reed and B. Simon, Functional Analysis, Academic Press (1980).

- [26] K. Fujisawa, M. Fukuda, K. Kobayashi, M. Kojima, K. Nakata, M. Nakata and M. Yamashita, *SDPA (SemiDefinite Programming Algorithm) and SDPA-GMP Users Manual Version 7.1.0*, Research Reports on Mathematical and Computing Sciences, B-448 (2008); M. Nakata, *A numerical evaluation of highly accurate multiple-precision arithmetic version of semidefinite programming solver:SDPA-GMP, -QD and -DD*, the proceedings of 2010 IEEE Multi-Conference on Systems and Control, 29-34, 2010.
- [27] J. Cimprič, J. math. anal. appl., 369, no. 2, str. pp. 443-452 (2010).
- [28] D. Petz, *An Invitation to the Algebra of Canonical Commutation Relations*, Leuven University Press, Leuven (1990).
- [29] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, North-Holland Publishing Company (1982).

Appendix

A Bounding the norm of g_c^r

The goal of this section is to prove the next lemma, from which Lemma 5 is a direct corollary.

Lemma 9. *Let $c > 2$. Then,*

$$l_1(g_c^r) \leq \frac{c}{c-2}. \quad (47)$$

The proof of this lemma, we will rely on the next two propositions.

Proposition 10.

$$a^k(a^k)^* = \sum_{m=0}^k \frac{k!^2}{m!^2(k-m)!} (a^m)^* a^m. \quad (48)$$

Proof. Clearly, $a^k(a^k)^* = \sum_{l,m} s_{l,m} (a^*)^l a^m$. Evaluating the mean value of the Schrödinger representation of both polynomials with respect to an arbitrary coherent state $|\alpha\rangle$, we have that

$$\sum_{l,m} s_{l,m} (\alpha^*)^l \alpha^m = e^{-|\alpha|^2} \sum_{j=0}^{\infty} \frac{(j+k)!}{j!^2} |\alpha|^{2j} =: g(|\alpha|^2). \quad (49)$$

It follows that

$$s_{l,m} = \delta_{l,m} \frac{1}{m!} \left. \frac{d^m g(x)}{dx^m} \right|_{x=0}. \quad (50)$$

Now, it can be proven, by induction, that

$$\frac{d^m g(x)}{dx^m} = \frac{k!}{(k-m)!} e^{-x} \sum_{j=0}^{\infty} \frac{(j+k)!}{j!(j+m)!} x^j. \quad (51)$$

if $m \leq k$, while $\frac{d^m g(x)}{dx^m} = 0$ if $m > k$. The statement of the proposition follows from these two relations. \square

Proposition 11. Let $\#\{\bar{t} \in \mathbb{N}^n : \|\bar{t}\|_1 = k\}$ denote the number of elements of \mathbb{N}^n satisfying $\|\bar{t}\|_1 = k$. Then,

$$\#\{\bar{t} \in \mathbb{N}^n : \|\bar{t}\|_1 = k\} = \frac{(n+k-1)!}{(n-1)!k!}. \quad (52)$$

Proof. Our aim is to compute the number of ways in which k identical balls can be contained in n different boxes. Clearly, any possible configuration can be represented uniquely by a sequence of k dots “.” and $n-1$ bars “|”. The number of balls n_1 in box 1 would then correspond to the number of dots on the left of the first bar; the number of balls n_j inside box j , for $2 \leq j \leq n-1$, to the number of dots between the $j-1^{th}$ and the j^{th} bars; the number of balls in box n , to the number of dots on the right of the $n-1^{th}$ box. For instance, the configuration $n_1 = 2, n_2 = 0, n_3 = 1$ would be represented by “..|.”.

It is elementary that the number of permutations of $n+k-1$ elements, out of which $n-1$ and k are indistinguishable, is equal to $\frac{(n+k-1)!}{(n-1)!k!}$. \square

Proof of Lemma 9. Proposition 10 implies that

$$l_1(a^k(a^k)^*) = \sum_{m=0}^k \frac{k!^2}{m!^2(k-m)!} \leq \sum_{m=0}^k \frac{k!}{m!(k-m)!} k! = 2^k k!. \quad (53)$$

The last expression is logarithmically superadditive, i.e., $\prod_i 2^{k_i} k_i! \leq 2^k k!$, for all sets of natural numbers $\{k_1, k_2, \dots\}$ such that $\sum_i k_i = k$. It follows that the bound given by eq. (53) also holds for $l_1(a^{\bar{t}}(a^{\bar{t}})^*)$, with $\bar{t} \in \mathbb{N}^n$, $\|\bar{t}\|_1 = k$. We thus have that

$$\begin{aligned} l_1(g_c^r) &\leq \sum_{|\bar{t}| \leq r} \frac{(n-1)!}{c^{|\bar{t}|_1} (n + \|\bar{t}\|_1 - 1)!} l_1(a^{\bar{t}}(a^{\bar{t}})^*) \leq \sum_{k=0}^r \frac{\#\{\bar{t} : \|\bar{t}\|_1 \leq k\}}{\binom{n+k-1}{n-1}} \frac{2^k}{c^k} = \\ &= \sum_{k=0}^r \left(\frac{2}{c}\right)^k \leq \sum_{k=0}^{\infty} \left(\frac{2}{c}\right)^k = \frac{c}{c-2}, \end{aligned} \quad (54)$$

where in the third inequality we have made use of Proposition 11. \square

B Anti-normal ordered monomials

The following appendix establishes Lemma 7.

Proposition 12. *For any $k \in \mathbb{N}$,*

$$a^{k+1}(a^{k+1})^* - a^*a^k(a^k)^*a \in \Sigma^2. \quad (55)$$

Proof. Using the CCRs, we have that

$$a^*a^k(a^k)^*a = -ka^{k-1}(a^k)^*a - (k+1)a^k(a^*)^k + a^{k+1}(a^*)^{k+1}. \quad (56)$$

Now, by induction, we have that, for any $0 \leq l \leq k$, $a^l(a^k)^*a^{k-l} \in \Sigma^2$. Indeed, for $l = 0$ the result is obvious. Suppose now that the result holds for l . Then

$$a^{l+1}(a^k)^*a^{k-l-1} = a^l a(a^k)^*a^{k-l-1} = ka^l(a^*)^{k-1}a^{k-l-1} + a^l(a^k)^*a^{k-l}, \quad (57)$$

and the last expression belongs to Σ^2 by hypothesis.

It follows that

$$a^{k+1}(a^{k+1})^* - a^*a^k(a^k)^*aa = (k+1)a^k(a^k)^* + ka^{k-1}(a^k)^*a \in \Sigma^2. \quad (58)$$

□

Proposition 13. *Let $s \in \mathcal{W}_1$ be an arbitrary monomial of length k . Then,*

$$a^k(a^k)^* - ss^* \in \Sigma^2. \quad (59)$$

Proof. We will prove the proposition by induction. Suppose, thus, that the proposition holds for all monomials of length smaller or equal than k , and let s be an arbitrary monomial with $|s| = k + 1$. There are two possibilities:

1. $s = a\tilde{s}$, with $|\tilde{s}| = k$. Then we have that

$$a^{k+1}(a^{k+1})^* - ss^* = a(a^k(a^k)^* - \tilde{s}\tilde{s}^*)a^* = \sum_i af_i f_i^* a^* \in \Sigma^2. \quad (60)$$

2. $s = a^*\tilde{s}$, with $|\tilde{s}| = k$. Then we have that

$$a^{k+1}(a^{k+1})^* - ss^* = \{a^{k+1}(a^{k+1})^* - a^*a^k(a^k)^*a\} + \{a^*(a^k(a^k)^* - \tilde{s}\tilde{s}^*)a\}. \quad (61)$$

The first term between brackets is a SOS by Proposition 12; the second term is a SOS due to the induction hypothesis.

To complete the induction we also have to show that the proposition also holds for $k = 1$. But this is trivial, since, in that case, $aa^* - ss^*$ equals 0 (1), for $s = a$ ($s = a^*$).

□

Lemma 7. *Let $s \in \mathcal{W}_n$ be a monomial. Then,*

$$\dagger ss^* \dagger - ss^* \in \Sigma^2. \quad (62)$$

Proof. Proposition 13 already shows that the lemma holds for $n = 1$. Now, suppose that the lemma holds for n , and let $s = tu \in \mathcal{W}_{n+1}$, with t (u) being a word with the letters $a_1, \dots, a_n, a_1^*, \dots, a_n^*$ (a_{n+1}, a_{n+1}^*). Let $\bar{t} \in \mathbb{N}^n$ be such that $\dagger t t^* \dagger = a^{\bar{t}}(a^{\bar{t}})^*$. Then,

$$\dagger ss^* \dagger - ss^* = a^{\bar{t}} a_{n+1}^{|u|} (a_{n+1}^{|u|})^* (a^{\bar{t}})^* - a^{\bar{t}} u u^* (a^{\bar{t}})^* + u \dagger t t^* \dagger u^* - u t t^* u^*. \quad (63)$$

The first two terms on the right hand side admit a SOS decomposition due to Proposition 13. The two remaining terms belong to Σ^2 because of the induction hypothesis. \square

C States in the Schrödinger representation

In this appendix, we demonstrate the following lemma.

Lemma 8. *Let L be a linear functional in \mathcal{W}_n . If $L(h^*h) \geq 0$ for any $h \in \mathcal{W}_n$ and there exist $c, d > 0, k \in \mathbb{N}$ such that for any monomial $s \in \mathcal{W}_n$, the relation*

$$|L(s)| \leq dc^{|s|} \Gamma\left(\frac{|s| + k + 1}{2}\right) \quad (64)$$

holds, then there exists a non normalized quantum state ρ (a non-negative trace class operator) such that

$$L(h) = \text{tr}(\rho \pi(h)), \quad (65)$$

for any polynomial $h \in \mathcal{W}_n$.

Proof. Suppose that, indeed, such a functional exists and define its characteristic function $\chi(\bar{\xi})$ as

$$\chi(\bar{\xi}) = \sum_{l=0}^{\infty} \frac{L((i\bar{\xi} \cdot \sigma \bar{R})^l)}{l!} =: L(e^{i\bar{\xi} \sigma \bar{R}}), \quad (66)$$

where $\bar{R} \in \mathcal{W}_n^{2n}$ is the vector of polynomials

$$\bar{R} = \left(\frac{a_1 + a_1^*}{\sqrt{2}}, \frac{a_1 - a_1^*}{i\sqrt{2}}, \dots, \frac{a_n + a_n^*}{\sqrt{2}}, \frac{a_n - a_n^*}{i\sqrt{2}} \right), \quad (67)$$

and σ denotes the symplectic form, i.e., $\sigma = \oplus_{l=1}^n \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Using (64), we have that

$$|\chi(\bar{\xi})| \leq d \sum_{l=0}^{\infty} \frac{(2n\hat{\xi}c)^l \Gamma(\frac{l+k+1}{2})}{l!}, \quad (68)$$

with $\hat{\xi} = \frac{1}{\sqrt{2}} \max_k \{|\xi_{2k+1} + i\xi_{2k+2}|\}$. That the last series converges for any value of $\hat{\xi}$ follows from the relation

$$\int_{-\infty}^{\infty} |x|^k e^{-x^2} dx = \Gamma\left(\frac{k+1}{2}\right). \quad (69)$$

This allows us to write

$$\sum_{l=0}^{\infty} \frac{a^l \Gamma\left(\frac{l+k+1}{2}\right)}{l!} = \int_{-\infty}^{\infty} e^{a|x|} |x|^k e^{-x^2} dx, \quad (70)$$

and the last integral converges for all $a \in \mathbb{R}$.

Now, define the operator

$$\rho \equiv \frac{1}{(2\pi)^n} \int d\bar{\xi} \chi(\bar{\xi}) W_{-\bar{\xi}}, \quad (71)$$

with $W_{\bar{\xi}} = e^{i\bar{\xi}\sigma\bar{R}}$ being the so called *Weyl operator* [28].

We will next prove that, for any polynomial h , $\text{tr}(\rho\pi(h)) = L(h)$.

Because $\text{tr}(W_{\bar{\xi}}W_{\bar{\eta}}) = (2\pi)^n \delta(\bar{\xi} + \bar{\eta})$, it follows that $\text{tr}(\rho W_{\bar{\xi}}) = \chi(\bar{\xi}) = L(W_{\bar{\xi}})$. Moreover, from the *Weyl relations*

$$W_{\bar{\xi}}W_{\bar{\eta}} = e^{-i\bar{\xi}\sigma\bar{\eta}/2} W_{\bar{\xi}+\bar{\eta}}, \quad (72)$$

it is immediate that $\text{tr}(\rho W_{\bar{\xi}}W_{\bar{\eta}}) = L(e^{-i\bar{\xi}\sigma\bar{\eta}/2} W_{\bar{\xi}+\bar{\eta}}) \equiv f(\bar{\xi}, \bar{\eta})$. We will now show that

$$f(\bar{\xi}, \bar{\eta}) = \lim_{r \rightarrow \infty} f^r(\bar{\xi}, \bar{\eta}), \quad (73)$$

where $f^r(\bar{\xi}, \bar{\eta}) \equiv \sum_{l,m=0}^r L\left(\frac{(i\bar{\xi}\cdot\sigma\bar{R})^l}{l!} \frac{(i\bar{\eta}\cdot\sigma\bar{R})^m}{m!}\right)$.

Note that f is analytic, and that f^r corresponds to a sort of truncated Taylor expansion of f with respect to the variables $\bar{\xi}, \bar{\eta}$, where only those monomials $\xi^{\bar{s}}\eta^{\bar{t}}$ with $\|\bar{s}\|_1, \|\bar{t}\|_1 \leq r$ are present. Indeed, if we take any number of derivatives of ξ_i, η_j (no more than r of each) on both sides of (72) and then evaluate on the point $\bar{\xi} = \bar{\eta} = 0$, we will arrive at two polynomials $p_1, p_2 \in W_n$ on each side. Since this is a general relation between operators, both polynomials must be the same modulo the canonical commutation relations. It follows that

$$\frac{\partial^m \partial^{m'}}{\prod_{k=1}^m \partial \xi_{i_k} \prod_{l=1}^{m'} \partial \eta_{j_l}} f(\bar{\xi}, \bar{\eta}) \big|_{\bar{\xi}=\bar{\eta}=0} = L(p_1) = L(p_2) = \frac{\partial^m \partial^{m'}}{\prod_{k=1}^m \partial \xi_{i_k} \prod_{l=1}^{m'} \partial \eta_{j_l}} f^r(\bar{\xi}, \bar{\eta}) \big|_{\bar{\xi}=\bar{\eta}=0}, \quad (74)$$

for $m, m' \leq r$. Since $f(\bar{\xi}, \bar{\eta})$ is analytic in $\bar{\xi}, \bar{\eta}$, the sum of the absolute values of the coefficients of its Taylor expansion must converge. This implies that $\lim_{r \rightarrow \infty} f(\bar{\xi}, \bar{\eta}) - f^r(\bar{\xi}, \bar{\eta}) \rightarrow 0$ for fixed $\bar{\xi}, \bar{\eta}$, and so we get equation (73).

Analogously, one can prove that

$$\text{tr}(\rho \prod_{j=1}^k W_{\bar{\xi}_j}) = \lim_{N \rightarrow \infty} L \left(\prod_{j=1}^k \sum_{l=0}^N \frac{(i\bar{\xi}_j \cdot \sigma \bar{R})^l}{l!} \right), \quad (75)$$

for any k .

Now, let $p = \prod_{k=1}^m p_k \in W_n$, with $p_k \in \{\frac{a_i + a_i^*}{2}, \frac{a_i - a_i^*}{2i}\}_{i=1}^n$. Then,

$$\frac{L(p)}{i^m} = \frac{\partial^m}{\partial \epsilon_1, \dots, \epsilon_k} L \left(\prod_{k=1}^m \sum_{l=0}^{\infty} \frac{(i\epsilon_k s_k)^l}{l!} \right) \Big|_{\bar{\epsilon}=0} = \frac{\partial^m}{\partial \epsilon_1, \dots, \epsilon_k} \text{tr}(\rho \prod_{k=1}^m e^{i\epsilon_k p_k}) = \frac{\text{tr}(\rho \pi(p))}{i^m}. \quad (76)$$

By taking linear combinations of the former expectation values, we thus have that $\text{tr}(\rho \pi(h)) = L(h)$ for any $h \in \mathcal{W}_n$.

It only rests to show that ρ is a quantum state, i.e., it is trace-class and positive semidefinite.

From the quantum Bochner-Kinchin theorem [29], we know that $\chi(\bar{\xi})$ is the characteristic function of a non normalized quantum state if and only if

- i) $\chi(\bar{\xi})$ is continuous at the origin.
- ii) for any $r \in \mathbb{N}$, $\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_r \in \mathbb{R}^{2n}$, and $c_1, c_2, \dots, c_r \in \mathbb{C}$ the relation

$$\sum_{k,l=0}^r c_k c_l^* \chi(\bar{\xi}_k - \bar{\xi}_l) e^{i\bar{\xi}_k \sigma \bar{\xi}_l / 2} \geq 0 \quad (77)$$

holds.

From (68) we know that $\chi(\bar{\xi})$ is not only continuous everywhere but even analytic. On the other hand, from (73) we have that

$$\sum_{k,l=0}^r c_k c_l^* \chi(\bar{\xi}_k - \bar{\xi}_l) e^{i\bar{\xi}_k \sigma \bar{\xi}_l / 2} = \lim_{m \rightarrow \infty} L(h_m h_m^*), \quad (78)$$

where $h_m = \sum_k c_k \sum_{l=0}^m \frac{(i\bar{\xi}_k \cdot \sigma \bar{R})^l}{l!}$. Since $h_m \in \mathcal{W}_n$, $L(h_m h_m^*) \geq 0$, and so the above limit is non-negative.

□